K.Harrison
BNL, 7 May 2002

Overview of GANGA

– First steps towards GANGA
– Outline of required functionality
– Some existing technology
– Conclusions

First steps towards GANGA

– ATLAS and LHCb develop applications within a common framework: Gaudi/Athena
– Both collaborations aim to exploit potential of Grid for large-scale, data-intensive distributed computing

$\Rightarrow$ Simplify management of analysis and production jobs for end-user physicists by developing tool for accessing Grid services with built-in knowledge of how Gaudi/Athena works: Gaudi/Athena and Grid Alliance (GANGA)

– First ideas for GANGA presented by P.Mato in October 2001
– Development of GANGA encouraged by many people and supported in UK within GRIDPP project by PPARC

⇒ Two joint ATLAS/LHCb posts funded to end of 2004:

- ♦ K.Harrison (Cambridge)
  - ⇒ Background in experimental particle physics
    For LHC: worked briefly with ATLAS TileCal Group; spent two years as software coordinator for LHCb Muon Group
- ♦ A.Soroko (Oxford)
  - ⇒ Background mainly in theoretical solid-state physics
    For LHC: worked with ATLAS Muon Group

- GANGA development is very much at the beginning
- Current work focuses on understanding requirements and reviewing existing technology
- Aim to have a draft design by August, and a command set for basic job configuration and submission by October

Outline of required functionality

1) Tasks prior to job execution

– Perform job configuration: select algorithms to run and set properties; specify input event data, detector characteristics, requested output, etc

  ⇒ Common configurations could be stored in a database and retrieved using high-level commands

  ⇒ User would have possibility of modifying settings and storing personalised configurations in his/her own area

– Determine job requirements in terms of software products needed (executables, libraries, databases, etc)

– Query catalogues to find location of input data; decide whether to replicate

– Estimate necessary resources (CPU, storage, network, etc)

– Apply collaboration policies (quotas, privileges, etc) and carry out security checks (verify credentials)
– Translate user configuration and input request into one or more files of Job Description Language

$\Rightarrow$ Allow decomposition of single large job into several smaller jobs

– Choose computing element(s) according to some optimisation criteria and submit job(s)

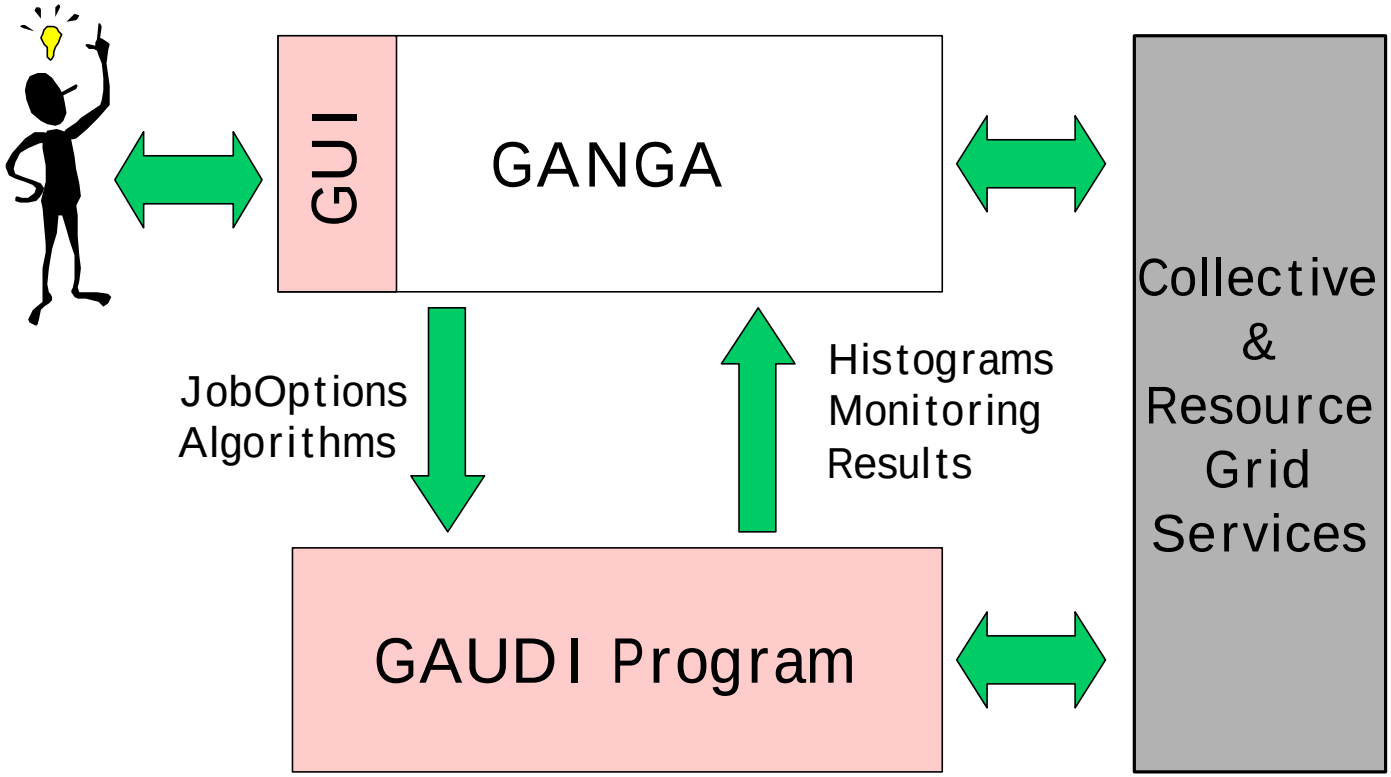
2) Tasks during execution

– Permit monitoring of job progress: display counter values, messages, histograms, etc
– Allow error recovery

3) Tasks after execution

– Collect output(s), merging where large job is decomposed prior to submission
– Where necessary, copy output(s) to backed-up mass-storage device
– Update relevant bookkeeping databases

GANGA
Gaudi/Athena and Grid Alliance



(From P.Mato)

Some existing technology

1) General-purpose Grid portals

– Grid portals not tied to a single application/framework are under development by several groups

⇒ Strategy for GANGA might be to take one of these general-purpose Grid portals and add functionality specific to Gaudi/Athena

⇒ Have started examining two candidates:

- ♦ Alice Environment (AliEn)
- ♦ Grid Enabled Web Environment for Site-Independent User Job Submission (GENIUS)

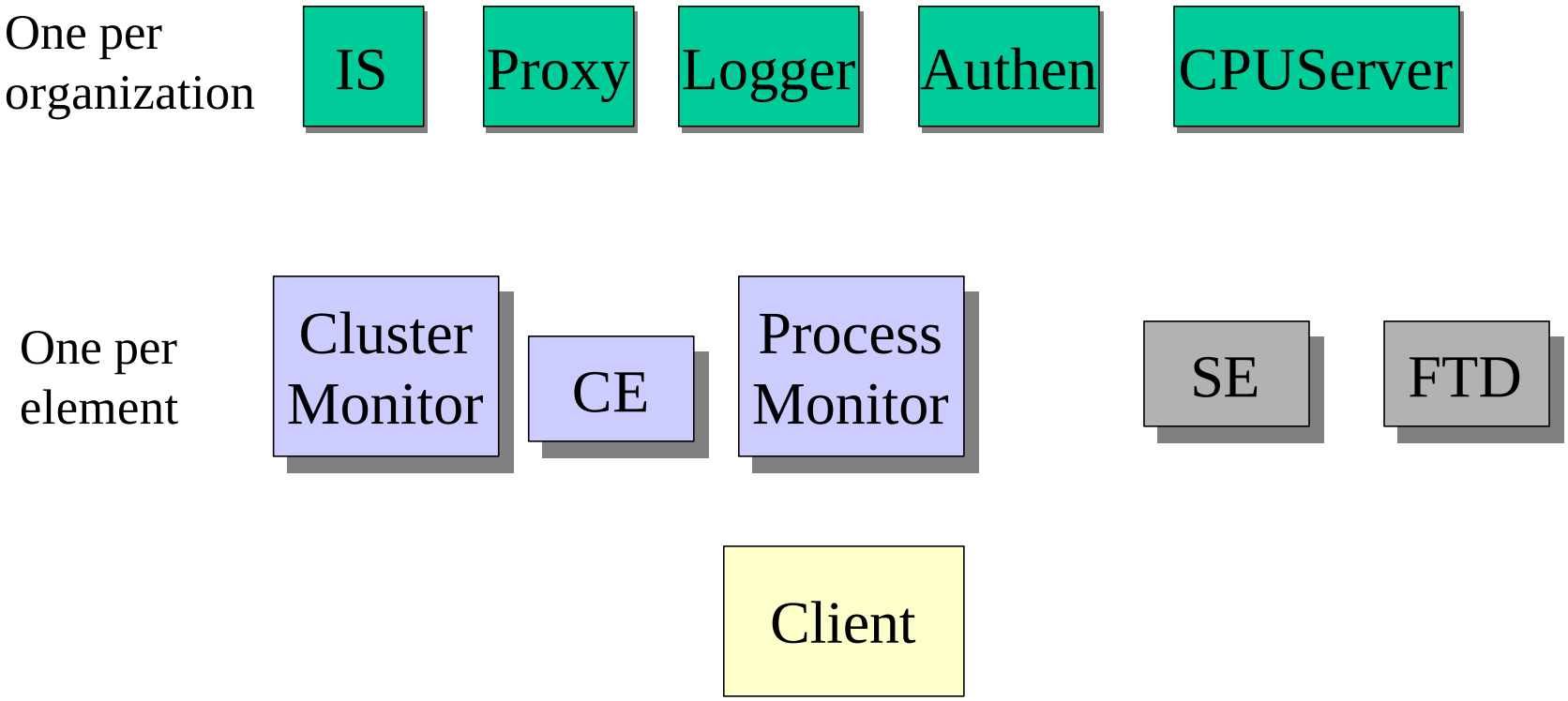
⇒ Guidance from P.Saiz (AliEn) and R.Barbera (GENIUS), both extremely helpful

AliEn

– Under development by Alice Offline Group, but not specific to Alice
– Uses iVDGL or EDG middleware, Globus toolkit, and a variety of external modules (SOAP, PAM, SWIG, etc); based on Perl
– User access via machine on which AliEn is installed:
  ♦ Command-line interface allows authentication, access to distributed catalogue, job submission, etc
  ♦ With appropriate module installed, also have GUI interface
– Web interface under development?

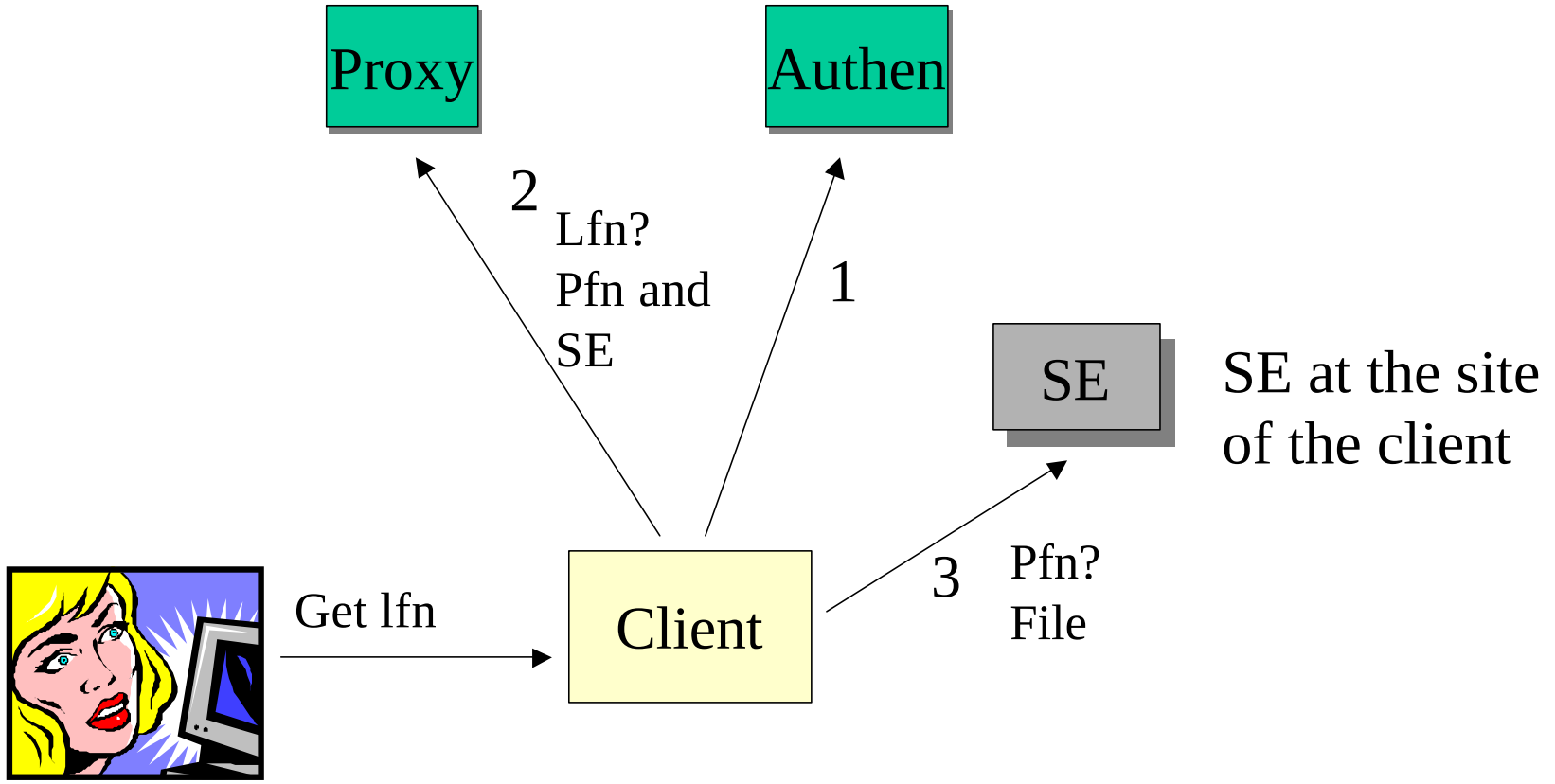
⇒ AliEn seems still at an experimental stage, at least for non-Alice users: some teething problems encountered, but has a lot of potential

AliEn Picture



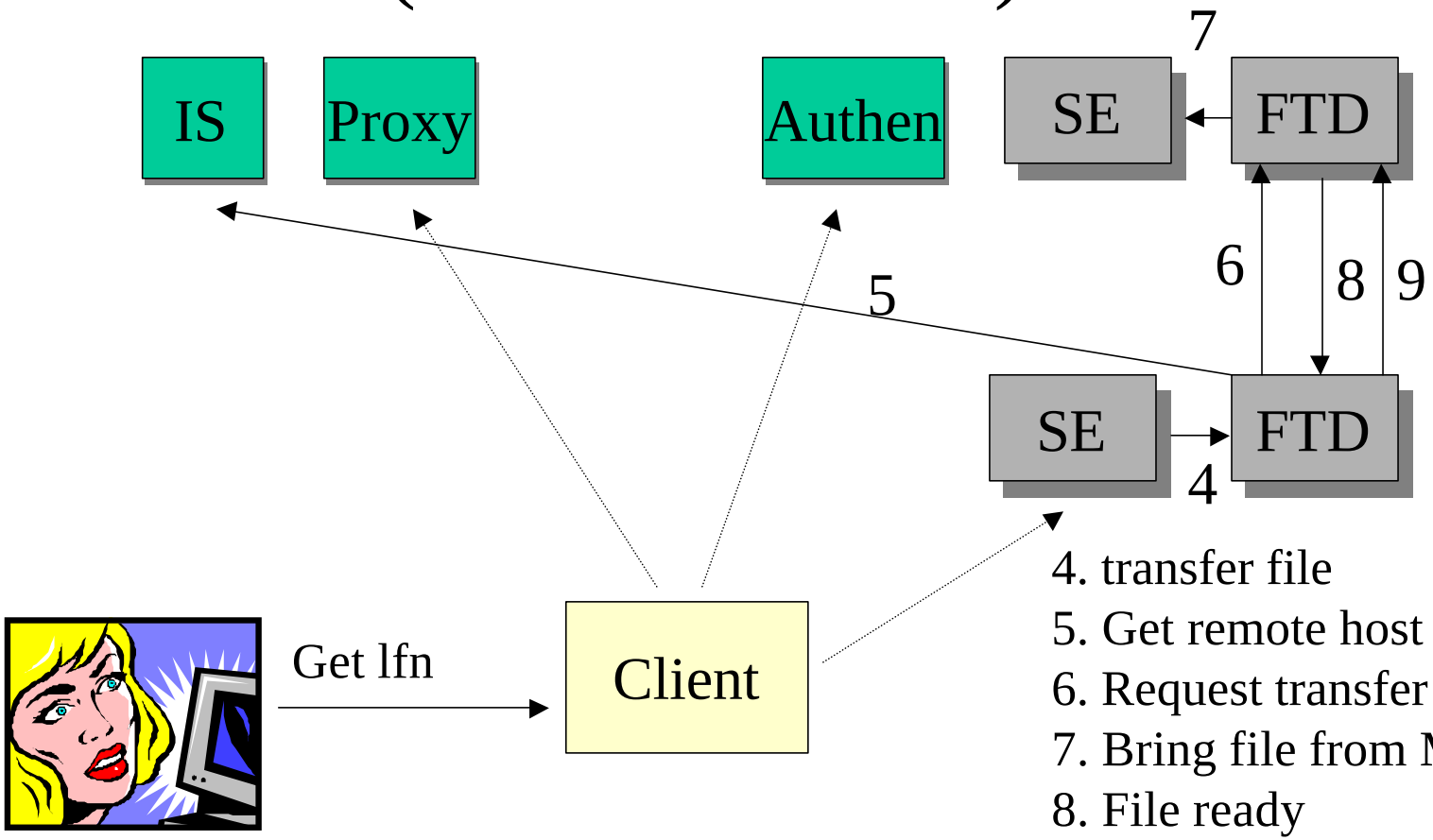
(From P.Saiz)

Getting a file (from local SE)



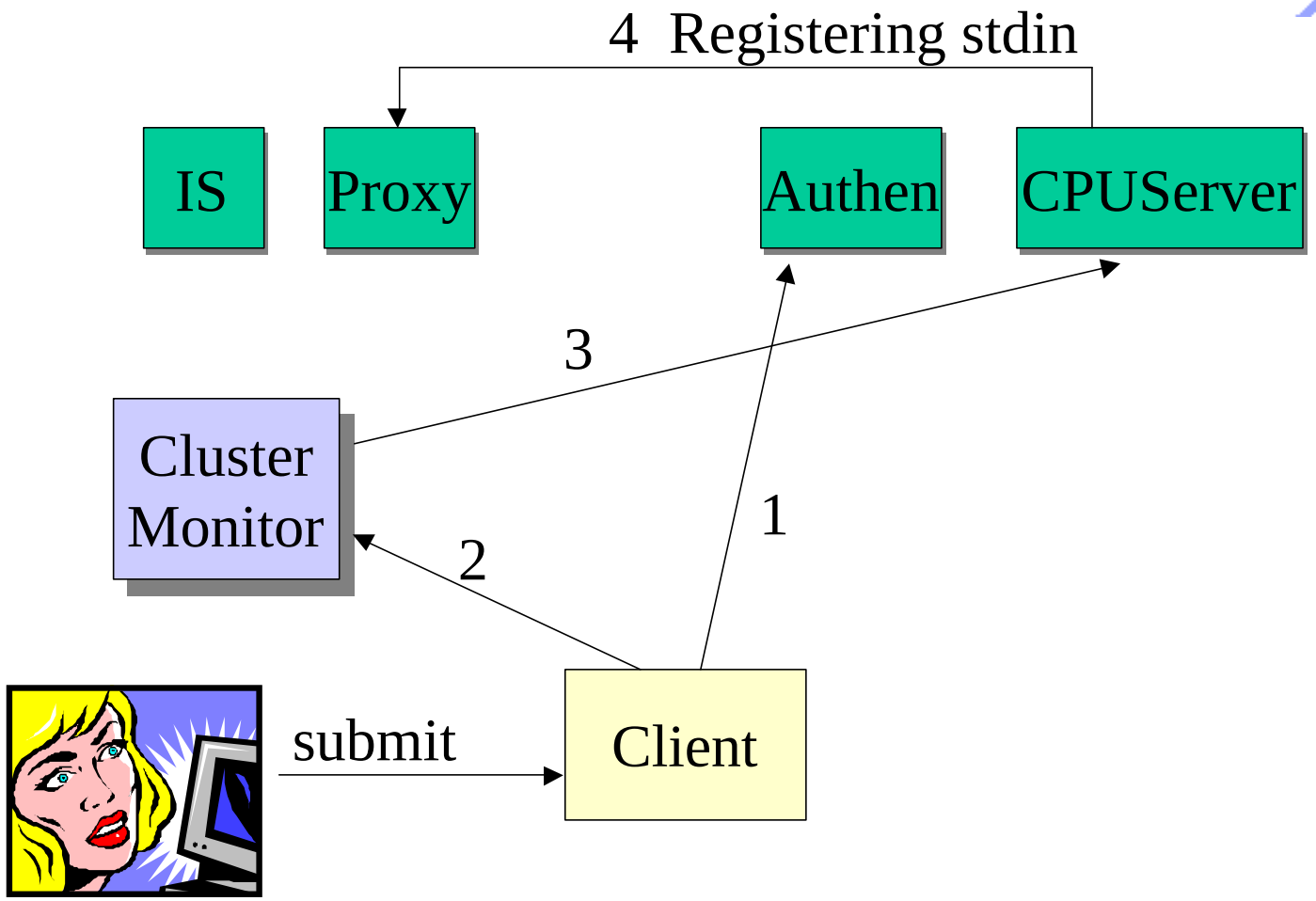
(From P.Saiz)

Getting a file (from remote SE)



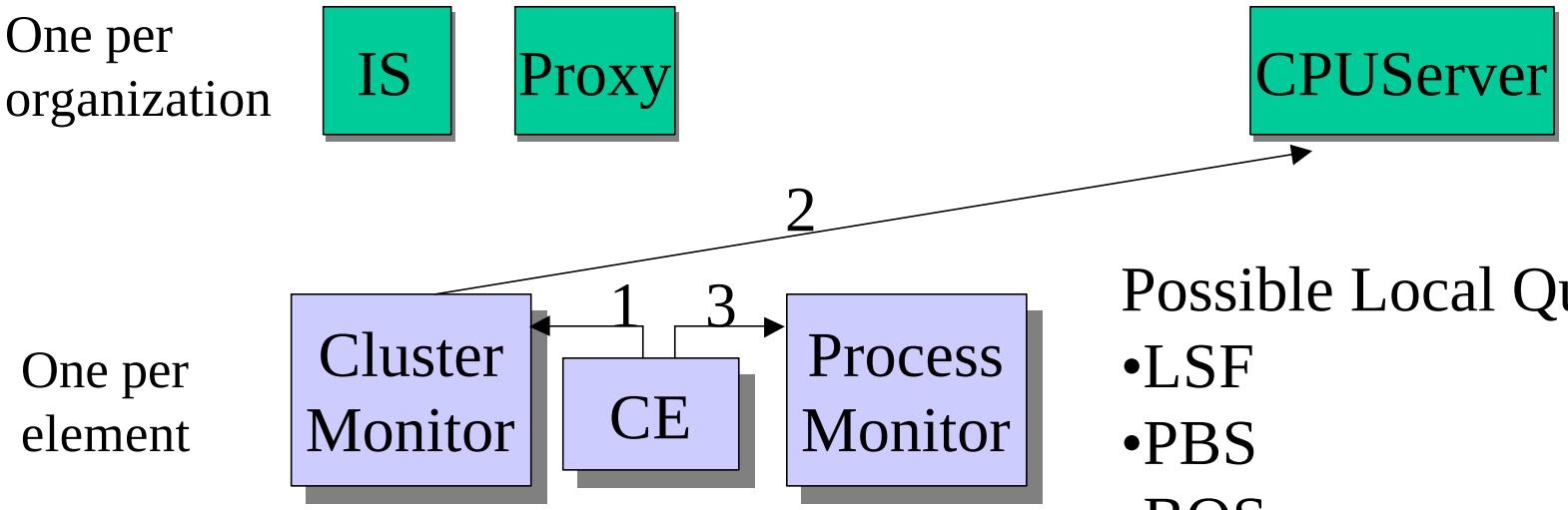
(From P.Saiz)

Submit a job



(From P.Saiz)

Executing a job



Possible Local Queues:

- LSF
- PBS
- BQS
- Globus
- CONDOR
- DQS

(From P.Saiz)

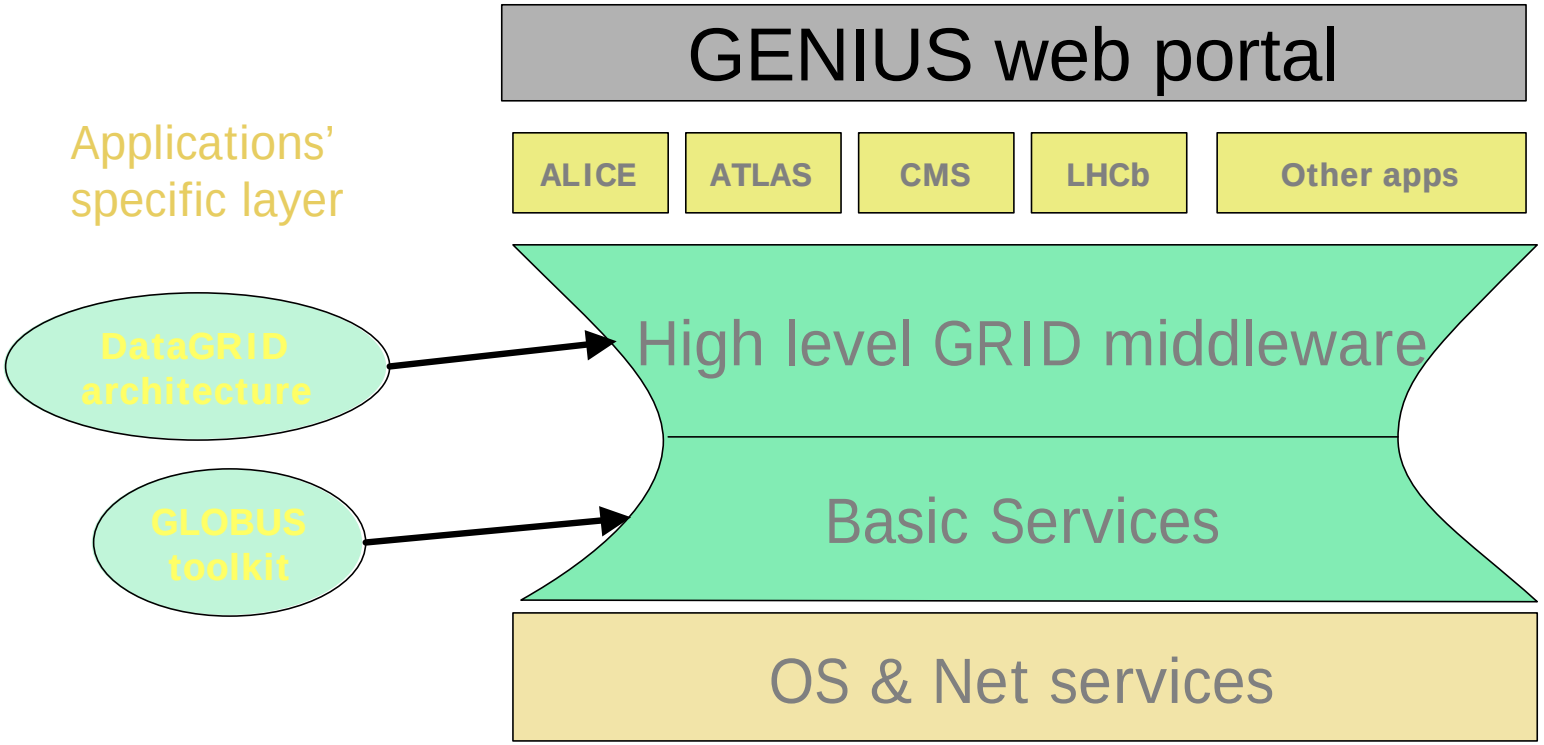
GENIUS

– Under development by INFN and NICE srl in framework of INFN Grid Project
– Uses EDG middleware, Globus toolkit and the EnginFrame framework of NICE srl; based on Java and XML
– User obtains account on interface machine where GENIUS is installed, uploads Grid certificates, then has Testbed access via web page from anywhere (desktop, laptop, PDA, WAP telephone,etc)
– Incorporates all services made available with first release of EDG middleware
– Allows files to be created/edited/deleted on interface machine; allows access to interface machine via web using VNC

⇒ Very impressive - elegant, straightforward and easy to use
⇒ Need to understand procedure for adding application-specific modules (Gaudi/Athena awareness)

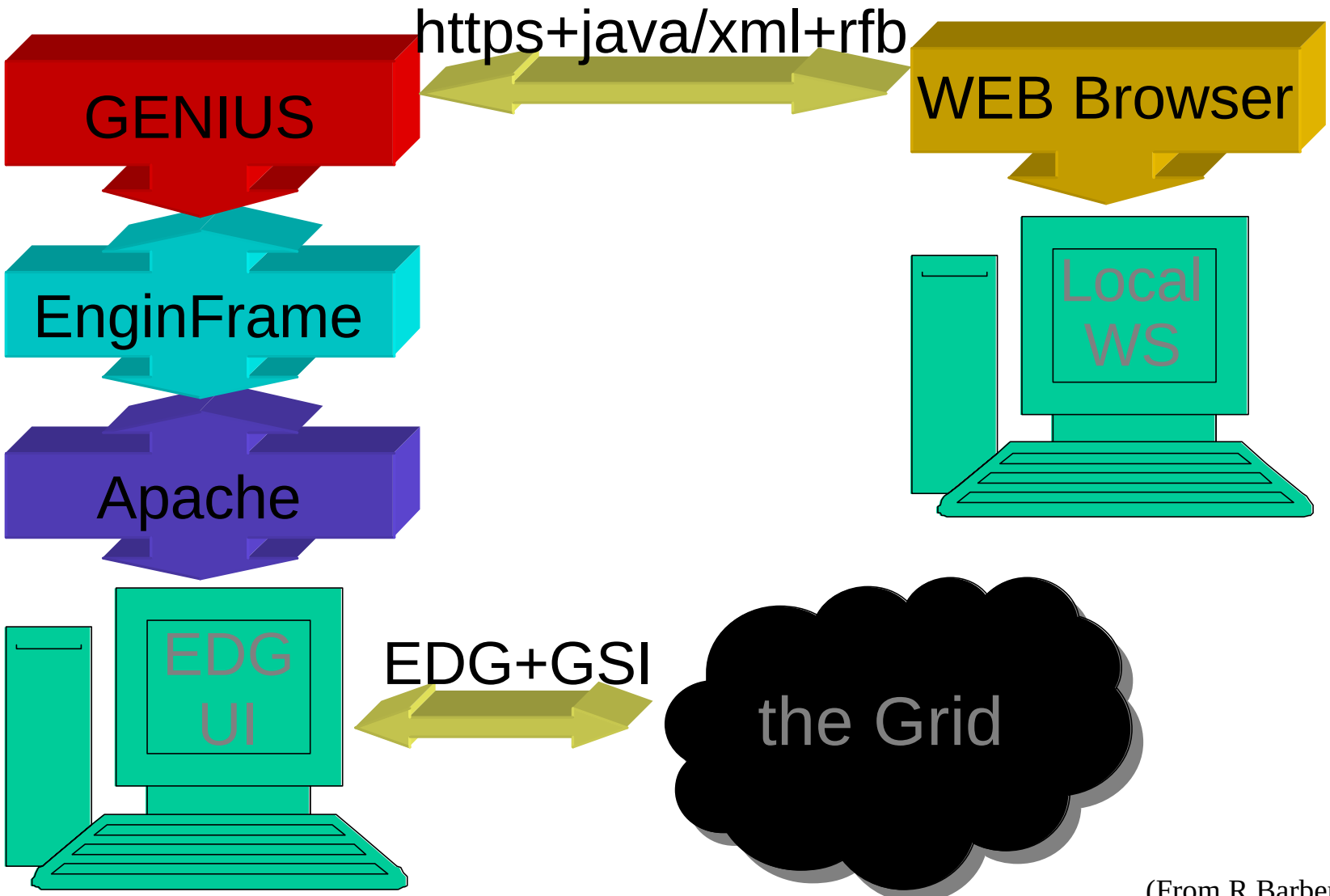
GENIUS®

(Grid Enabled web eNvironment for
site Independent User job Submission)



(From R.Barbera)

GENIUS: how it works



(From R.Barbera)

2) Production tools

– Production manager will be a special class of GANGA user (will typically have substantial resource requirements)

– Have started looking at Simulation for LHCb and its Integrated Control Environment (SLICE)

⇒ Help provided by E.vanHerwijnen

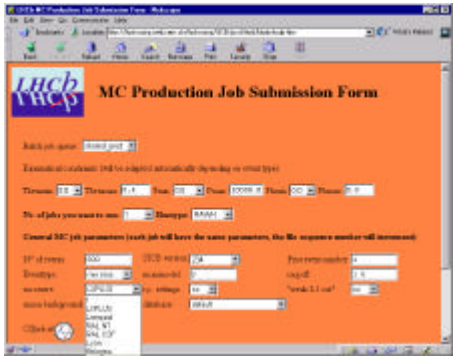
– In a non-Grid environment, SLICE performs several tasks similar to those foreseen for GANGA:

- ♦ Production requests to distributed facilities are submitted via a web page
- ♦ Java servelets create job scripts and options files
- ♦ Production is monitored using control system based on PVSS
- ♦ Update of bookkeeping database, transfer of output data to mass storage and quality checks performed automatically

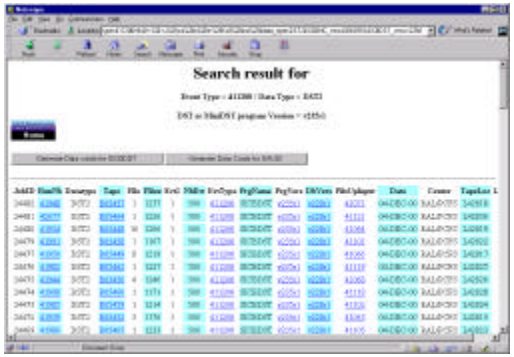
– Grid-based system at experimental stage

⇒ Provides valuable pointers for GANGA development

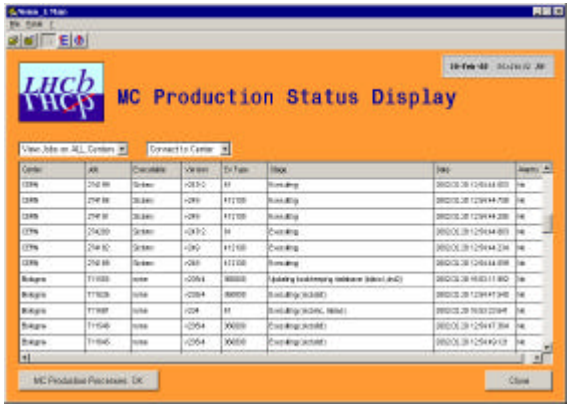
Submit jobs remotely view



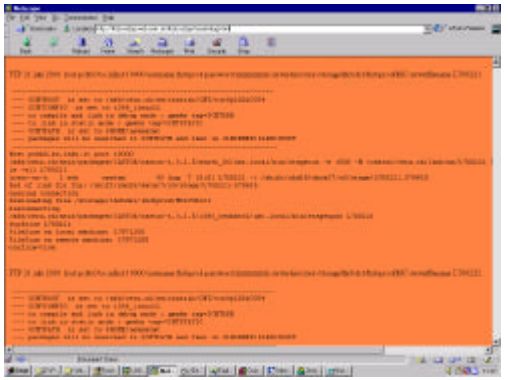
Update bookkeeping database



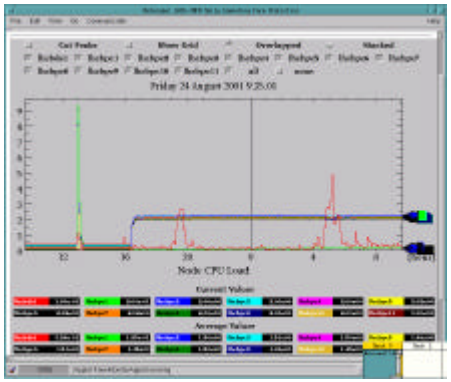
Execute on farm



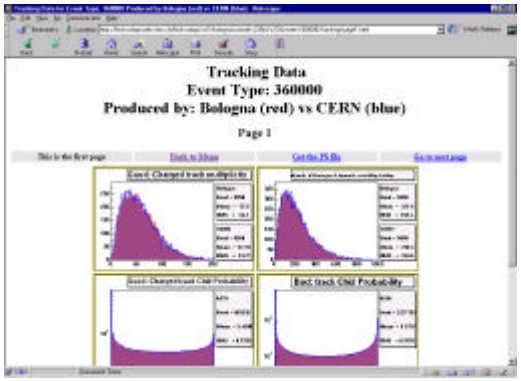
Transfer data to Mass store



Monitor performance of farm via Web



Data Quality Check



(From E.vanHerwijnen)

Conclusions

– Work on GANGA development has started
– Joint ATLAS/LHCb Grid Applications Meeting to be held 22-23 May at Cosener's House, Abingdon, UK

⇒ Expect to have discussion of GANGA requirements and existing technology

– General-purpose Grid portals are available (AliEn, GENIUS, others?) and can provide a starting point for GANGA
– Production tools set up for non-Grid environment are being studied as a source of ideas
– Aim to have a draft design for GANGA by August, and a command set for basic job configuration and submission by October

⇒ This is an open project - wants to take the best of what exists already, and involve all interested parties